



**Food and Agriculture
Organization of the United Nations**

**Evaluation Service
(PBEE)**

**DRAFT
AUTO-EVALUATION
GUIDELINES**

July 2007

*“The most important purpose of evaluation
is not to prove but to improve.”*

Egon Guba (writer on qualitative research methods)

Table of Content

INTRODUCTION.....	1
A. What is Auto-Evaluation in FAO?.....	1
B. Purpose of the Guidelines.....	1
C. The Role of the Evaluation Service.....	1
PART I: PROCEDURES.....	2
A. What to Evaluate: Scope	2
B. When to Plan and Perform an Auto-Evaluation?.....	2
C. Process and Roles – Who Does Auto-Evaluation?	2
PART II: PLANNING FOR AUTO EVALUATION.....	5
A. Identifying Evaluation Issues	5
B. Deciding on Evaluation Methodology – Participants and Techniques	6
C. Estimating a Budget.....	8
D. Drafting and Circulating Terms of Reference	8
PART III: MANAGING AN AUTO EVALUATION	10
A. Selecting and Sequencing Data Collection Techniques.....	10
B. Following-Up on Auto-Evaluation Recommendations	11
ANNEXES.....	12
Annex 1: Examples of Issues in PE Auto-Evaluations	13
Annex 2: Outline for Terms of Reference.....	14
Annex 3: Auto-Evaluation Report Outline	16
Annex 4: Desk Study.....	18
Annex 5: SWOT Analysis	19
Annex 6: Interviews and Focus Groups	20
Annex 7: Questionnaires	21
Annex 8: Web Statistics.....	22
Annex 9: Bibliography.....	23

List of Acronyms

ADG	Assistant Director-General
AE	Auto-Evaluation
CP	Continuing Programme Activity
CSO	Civil Society Organisation
DG	Director-General
DSA	Daily Subsistence Allowance
DOs	Decentralized Offices
KCEW	Knowledge Exchange & Capacity Building Division (Former GILW)
GPA	Gender Plan of Action
MTP	Medium Term Plan
NGO	Non-governmental Organization
PAIA	Priority Area for Inter-disciplinary Action
PBE	Office of Programme, Budget and Evaluation
PBEE	Evaluation Service
PE	Programme Entity
PIR	Programme Implementation Report
PIRES	Programme Planning, Implementation Reporting and Evaluation Support System
PWB	Programme of Work and Budget
SWOT	Strengths, Weaknesses, Opportunities and Threats
TOR	Terms of Reference
TP	Technical Project
TS	Technical Services agreement
WHO	World Health Organization

INTRODUCTION

A. What is Auto-Evaluation in FAO?

Auto- or self-evaluation in FAO is defined as: "*A decentralized and participatory process of rigorous review of programme achievements and outcomes by programme managers*"¹. FAO auto-evaluations are intended to strengthen the results-based management of the Organization for the purposes of improving programmes, enhancing organizational learning and providing a sound basis for informed decision-making.

B. Purpose of the Guidelines²

These *Auto-evaluation guidelines* are intended to:

- ▣ Introduce the auto-evaluation process by outlining the procedures, quality standards, planning and management of the auto-evaluation;
- ▣ Explain methodological approaches and highlight common methods; and
- ▣ Provide practical guidance to auto-evaluation teams.

C. The Role of the Evaluation Service

The Evaluation Service has been providing support to Programme staff and external consultants, since the launch of auto-evaluation (AE) in FAO in 2004. The role of the Evaluation Service in the context of AEs continues to be in:

- ▣ Providing general advice and consultation (e.g. preparation, review and clearance of Terms of Reference (TOR), checking adherence to timing and quality standards, coverage analysis, commenting on draft reports);
- ▣ Providing methodological advice and support (e.g. training and advice on methods and analysis, such as questionnaire design and analysis, web traffic analysis; facilitation of brainstorming meetings, focus group interviews and SWOT sessions);
- ▣ Assisting in selection of external consultants;
- ▣ Allocating financial support for AEs; and
- ▣ Following up on past AEs (e.g. summaries of AEs for the Programme Committee, and starting in 2007- Management Response to AE recommendations).

¹ From the "Guiding Principles for Pre-Evaluation Monitoring, Annual Assessment and Periodic Auto-evaluation of Technical and Economic Programmes", available on: http://www.fao.org/docs/eims/upload/215035/78491_en_eval_Guiding_Principles.doc.

² Users are encouraged to consult the PBEE website for information on previous auto-evaluations and current developments.

PART I: PROCEDURES

A. What to Evaluate: Scope

All Technical Projects (TPs) and Continuing Programme activities (CPs) should be auto-evaluated during a six-year period. Priority Areas for Inter-disciplinary Action (PAIAs) and Decentralized Offices (DOs) may also undergo an auto-evaluation³. The Programme Committee endorsed the proposal that AEs become part of the Regular Programme in 2005. It is therefore mandatory that AEs be planned by all organizational units as part of their Regular Programme of Work, and be funded from the Regular Programme⁴.

Auto-evaluations will come in different shapes and sizes. The AE of one single Programme entity (PE) represents the minimum scope. It is not permitted to auto-evaluate one single major output. In some instances, several PEs can be clustered in one auto-evaluation for the sake of cost-effectiveness. The maximum scope is that of a few related PEs, or conversely, of an entire Programme. However, the larger the evaluation, the more difficult it is to plan for and coordinate.

Given the reformulation of PE structure in the Organization (2006-2007), the following also applies:

- ▣ Old Programme structure should feed into new Programme structure – i.e. parts of the underlying structure of a single *new* PE which is to be auto-evaluated may have been auto-evaluated. The findings of these preceding AEs can be useful in providing baseline data and perspectives for the AE of the new PE.

B. When to Plan and Perform an Auto-Evaluation?

Given that PBEE has the organizational responsibility for ensuring that all PEs are subject to systematic auto-evaluation over a six-year period, PBEE proposes a list of PEs⁵ to undergo auto-evaluation each year. In preparing the proposal, all previous evaluations of Programmes and PEs, as well as the timeliness of the proposed PEs are taken into consideration by PBEE. The auto-evaluation plan is then finalized in discussions between PBEE staff and Programme Co-ordinators, and followed by an ADG response to the PBEE proposal.

For technical PEs, an auto-evaluation may take place in the middle or near the end of their six-year cycle. Cluster auto-evaluations may call for some flexibility as some TPs may be not far advanced.

C. Process and Roles – Who Does Auto-Evaluation?

1. Oversight and Management Roles in Technical Departments

³ Key AE principles outlined in these guidelines also apply to AEs other than TPs and CPs.

⁴ Financial support continues to be made available by PBEE, to support selected auto-evaluations and facilitate the auto-evaluation process in the Organization.

⁵ Alternative PEs suggested by Divisions to PBEE will be considered, provided sufficient explanation is given.

The auto-evaluation process is intended to be truly participatory, involving at key junctures all the concerned technical units at Headquarters and in those DOs that have contributed substantially to the work in question⁶.

Responsibilities for oversight and management depend on the scope of each specific auto-evaluation. For cluster AEs, ADGs would normally retain oversight responsibility and nominate, among the Service Chiefs concerned, an overall Auto-evaluation manager for the coordination of the exercise. In contrast, the auto-evaluation of a single PE could be overseen by the concerned Division Director and managed by a Service Chief or Senior staff member.

All AE reports shall be copied to PBEE for comments and quality control. Once they are cleared by the relevant Division Director, they will be submitted to the ADG of the Department for approval. Regional Representatives will be copied AE reports for those PEs to which their Offices have contributed. The ADG will then review the report, taking into account comments received from PBEE and DOs.

Following the ADG's review, a final version will be produced by the AE manager and forwarded to all concerned, including DOs and PBEE, together with a Management Response (see Part III, Table 4) by the ADG. The Management Response will define senior management's position vis-à-vis the recommendations made in the AE report and address major issues identified.

2. External Inputs

An auto-evaluation will always involve some external expertise and inputs, e.g. (a) external consultants to facilitate and/or moderate the AE process, and (b) may also include a review of the AE report by an external peer group.

The main rationale behind external inputs is to combine the advantage of AE in terms of ownership of the results with an objective and independent judgement. External consultants will be briefed and guided by PBEE as well as the AE manager. Table 1 illustrates the auto-evaluation process with the stages, tasks and inputs of various actors involved.

⁶ Regional Representatives will be invited to provide inputs to auto-evaluation processes that concern their office.

Table 1: Overview of the Auto-Evaluation Process

AE Stage	Tasks	Programme Staff	External Consultant	PBEE
AE planning and timing	<ul style="list-style-type: none"> ▪ Selection of PEs/clusters to be evaluated ▪ Nomination of AE managers ▪ TOR preparation 	<ul style="list-style-type: none"> ▪ The ADG selects PEs/clusters to be evaluated in consultation with Division Directors, and approves selection of AE managers ▪ Evaluation issues collected through a brainstorming meeting ▪ TOR circulated to all concerned staff for comments 		<ul style="list-style-type: none"> ▪ Proposes entities to be auto-evaluated (memorandum to ADGs) ▪ General advice and consultation on auto-evaluations ▪ Assistance in TOR preparation ▪ Finalization: TOR submitted to PBEE for review and clearance
Establishing the organization of the AE	<ul style="list-style-type: none"> ▪ AE manager allocates evaluation-related work to staff ▪ Hiring external consultant(s) 	<ul style="list-style-type: none"> ▪ The AE manager and selected Programme staff form core AE team: identify external consultant(s), initiate desk studies, fine-tune the methodology, etc. 	<ul style="list-style-type: none"> ▪ Recruitment 	<ul style="list-style-type: none"> ▪ Allocation of PBE budget contribution to Divisions ▪ Assistance in finding external consultant(s) ▪ Briefing external consultant(s)
Conduct of the evaluation	<ul style="list-style-type: none"> ▪ Collection and analysis of evaluation material and data by the AE team 	<ul style="list-style-type: none"> ▪ SWOT exercise with (normally) HQ staff ▪ Telephone interviews and email contacts with staff in DOs ▪ All Programme staff are key source of information – to be collected through individual interviews, focus group interviews and/or email 	<ul style="list-style-type: none"> ▪ Preparation of methodological instruments (e.g. questionnaires, focus group guides) ▪ Finalize/review desk study ▪ Collection of data through various techniques employed ▪ Analysis of data and feedback 	<ul style="list-style-type: none"> ▪ Methodological advice and support (e.g. questionnaire design, facilitation, web traffic analysis)
First draft of the evaluation report	<ul style="list-style-type: none"> ▪ Report preparation ▪ Report submitted to ADG 	<ul style="list-style-type: none"> ▪ Draft discussed with, and commented upon by, all concerned staff including in concerned DOs ▪ Division Director (in most cases) clears draft for review by ADG 	<ul style="list-style-type: none"> ▪ Prepares first draft of AE report for circulation to AE team (with a copy to PBEE and concerned DOs) 	<ul style="list-style-type: none"> ▪ Comments on draft report, quality control
Preparation of the final report	<ul style="list-style-type: none"> ▪ Finalization of AE report incorporating comments made during review process 	<ul style="list-style-type: none"> ▪ AE manager guides report finalization ▪ ADG's office to prepare Management Response ▪ All concerned staff and DOs receive a copy of the final report 	<ul style="list-style-type: none"> ▪ Consolidates feedback from the AE team and PBEE ▪ Prepares final AE report 	<ul style="list-style-type: none"> ▪ Follows up on, and collects final AE reports ▪ Disseminates AE findings (e.g. on PBEE website, AE summaries for the Programme Committee)

PART II: PLANNING FOR AUTO EVALUATION

A. Identifying Evaluation Issues

Identifying evaluation issues is an important exercise that helps to define TOR, and informs the evaluation approach (including the preparation of questionnaires, checklists for interviews, etc). AE goes much beyond a passive description of activities and outputs, and should review PE *design and relevance, implementation constraints and opportunities, outputs, outcomes and achievements* against the PE objective(s) from a qualitative and quantitative standpoint. Following are evaluation criteria to be considered when conceptualizing evaluation issues:

- ▣ **Organizational relevance:** conformity to the Organization's mandate; relevance to strategic objectives and use made of FAO comparative advantages;
- ▣ **Stakeholder relevance:** relevance to the needs of users (e.g. government officials, policy makers, researchers, programme planners and development practitioners);
- ▣ **PE design:** rationale, coherence and clarity of PE design, including logical relationship between planned outputs and expected outcomes;
- ▣ **AE findings:** past AE (or evaluation) findings relevant to the PE currently being auto-evaluated should be revisited (if applicable);
- ▣ **Partnerships:** strength and use made of internal and external partnerships;
- ▣ **Resources:** adequacy of staff and financial resources;
- ▣ **Performance:** overall performance in output production, particularly against qualitative and quantitative targets set out in the MTP, quality of management, cost-efficiency;
- ▣ **Output quality:** quality and adequacy of outputs produced as assessed by subject-area specialists and/or by actual or potential users;
- ▣ **Effectiveness:** effectiveness (including cost-effectiveness) in the realization of outcomes and objectives at the Programme Entity level,
- ▣ **Interdisciplinarity and Gender:** contribution to PAIAs and the Gender Plan of Action, particularly when such contributions were planned in the MTP;
- ▣ **Sustainability:** the extent to which the outcomes of the PE are likely to have lasting effects; and
- ▣ **Emerging issues:** are there any emerging issues, such as unforeseen problems or opportunities for future action that should be considered?

AE managers are expected to expand and elaborate on the issues outlined above to make them programme-specific through a participatory process, for instance in a series of AE planning and brainstorming meetings. Canvassing participating staff's expectations will help ensure they view the evaluation as relevant and useful. The following two "trigger" questions may be useful in this exercise:

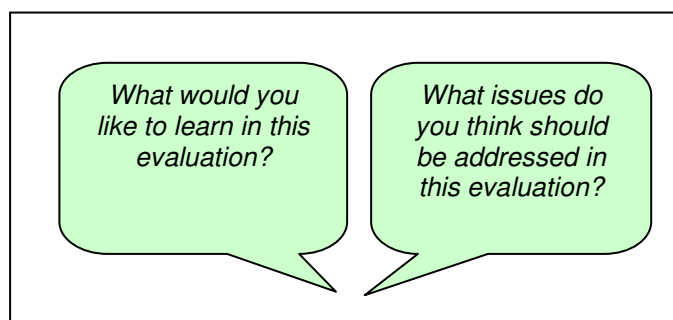


Diagram 1: Possible "trigger" questions to kick-start an AE brainstorming meeting

Note that consensus on what the key issues are and should be evaluated is, at times, difficult to reach. Some potential evaluation issues may prove to be contentious – however, the AE should have a wide enough scope to accommodate differing views also.

The AE consultations should produce a focused, synthetic (any overlapping issues should be clustered under a common evaluation theme), and relevant list of issues. As AEs are expected to follow a similar approach and yield comparable results, issues can be categorized under different headings (Design Issues, Implementation and Process Issues, Output Issues, Outcome Issues, Objective-level Issues, Cross-Sectoral Issues). As a rule of thumb, the list of evaluation issues should *never* exceed two pages. Annex 1 outlines an example of a list of evaluation issues.

B. Deciding on Evaluation Methodology – Participants and Techniques

1. Stakeholders and Participants

Once evaluation issues have been finalized, the next step is to identify the *sources* of information, including *who* the target participants will be (Diagram 2).

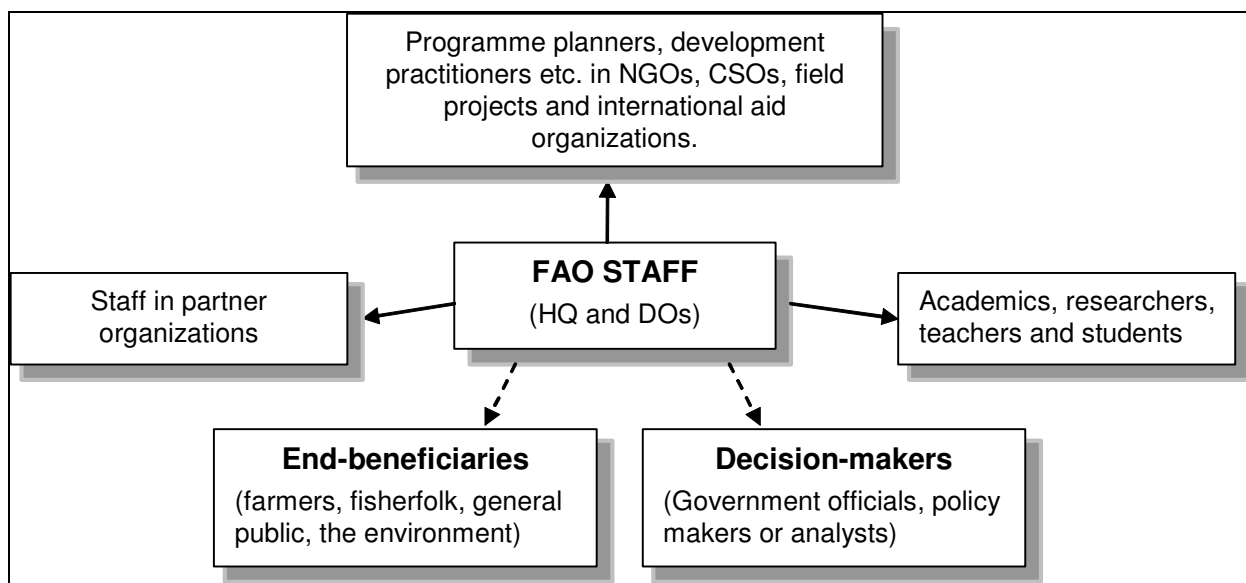


Diagram 2: Stakeholders and participants in auto-evaluations

FAO staff at Headquarters and DOs (if applicable to the area of work), are the primary informants that can give indications about users of their information and services. Based on past AE profiling⁷, FAO users of information products and services can be broadly categorized into three groups, according to the type of information sought from FAO:

- ▣ Government officials, policy makers or analysts;
- ▣ Academics, researchers, teachers and students; and
- ▣ Programme planners, development practitioners etc. in NGOs, CSOs, field projects and international aid organizations.

Decision-makers (e.g. Government officials, policy makers or analysts) are often difficult to reach in AEs; however, the possibility to include them may be explored (e.g. when decision-makers are also collaborators or implementing partners). End-beneficiaries are further along the chain of FAO users and are even more difficult to reach and cover in AEs.

2. Evaluation techniques

A desk study, together with Programme staff brainstorming sessions, is a good starting point in auto-evaluation. Once the AE has formally started, a SWOT analysis with Programme staff is useful to clarify the scope of the AE and identify common perceptions. Techniques used in AE include, but are not limited to the following:

- ▣ Interviews,
- ▣ Focus group interviews,

⁷ See document PC 93/4 b) titled: *Summary of Auto-Evaluations*, presented at the May 2005 Programme Committee

- ▣ Questionnaires, and

- ▣ Peer reviews.

A comparison of common auto-evaluation techniques in terms of advantages and disadvantages is outlined in Table 2 (see also Annexes 4-8). Note that it is common evaluation practice to use *triangulation*, or analysing an issue using several different methods (e.g. a survey and focus group interviews) and collecting information from different sources to check and validate evaluation results. In practice, this usually means using both qualitative (e.g. interviews and focus group interviews) and quantitative techniques (e.g. questionnaires).

C. Estimating a Budget

The next step in planning for an evaluation is for the AE manager to estimate how much resources are going to be required. The budget needs to include the actual cash budget needed for hiring consultants (including travel and DSA, if applicable), as well account for staff time put in by the auto-evaluation team⁸. Depending on the resources available, the budget may need to be revised (e.g. by deleting most costly data collection techniques or decreasing consultancy time) if selected evaluation methodology exceeds available resources. Note that cluster evaluations (combining several related PEs within a Department/Division) may be more cost-effective than reviewing each PE in isolation.

The methodology in the TOR should be outlined in sufficient detail to be able to match with reasonable precision the staff time and resources needed when estimating the budget (see Annex 2, section 7).

D. Drafting and Circulating Terms of Reference

By this stage, one should have all the elements for a draft TOR to be circulated to all concerned staff, and TOR should be cleared by the concerned ADG and by PBEE (Annex 2):

- ▣ A short background section,

- ▣ A list of issues to be evaluated (the auto-evaluation scope),

- ▣ A tentative methodology,

- ▣ A description of the persons/groups involved in the evaluation, and

- ▣ A budget estimate.

⁸ The total cost for AEs (including *both external and internal inputs*) will normally range from US\$20,000 and US\$50,000 depending on the size of the evaluated PE(s), or about 2 to 3 percent of the evaluated PE(s) budget.

Table 2: Advantages and drawbacks of some evaluation techniques

Evaluation Techniques	Advantages	Drawbacks
<p>Desk study</p> <ul style="list-style-type: none"> • Review of documentation (e.g. relevant past AE findings, back-to-office reports etc.) • MTP⁹ and other qualitative and quantitative indicators • Annotated bibliographies • Web traffic analysis (quantity of traffic on a particular website over time, publications trend analysis) 	<ul style="list-style-type: none"> ▪ Reduce the time needed for consultants or staff to access programme details or status of research ▪ Past AE findings may be relevant and a good starting point for the current AE ▪ Indicators are a strong accountability tool, since determined at planning stage; allows trends analysis ▪ Web traffic analysis is cheap and available (KCEW collects and publishes monthly web traffic) 	<ul style="list-style-type: none"> ▪ May take time to assemble, junior staff may need supervision, essential points may not be covered ▪ Indicators are difficult to verify, fail to capture unexpected developments, describe but do not explain ▪ Web traffic data misses vital information; geographic data biased towards developed countries access providers; quality of visits more important than quantity
<p>SWOT analysis</p> <ul style="list-style-type: none"> • Identify strengths, weaknesses, opportunities and threats in a qualitative way 	<ul style="list-style-type: none"> ▪ Participatory and transparent ▪ Good strategic tool: help focus on what is important and help define recommendations 	<ul style="list-style-type: none"> ▪ Best if all Programme staff involved ▪ Outcome partly depends on facilitation skills
<p>Interviews</p> <ul style="list-style-type: none"> • Usually semi-structured (basic list of issues to ask, with a possibility of exploring other related issues) • Collect qualitative information, e.g. quotes and “client stories” (narratives of user experiences) 	<ul style="list-style-type: none"> ▪ Help capture complex programmes aiming at varied outcomes ▪ Capture processes, complex issues and underlying reasons ▪ Help understand programme meaning to its stakeholders 	<ul style="list-style-type: none"> ▪ Takes time to interview and analyse, cannot be automated; interviewing skills needed ▪ Less good for programmes with repetitive, predetermined outcomes ▪ Some audiences distrust qualitative research, prefer statistics
<p>Focus group interviews</p> <ul style="list-style-type: none"> • Interactive group discussion lead by a facilitator (usually 6-12 participants), semi-structured • Participants are asked about their perceptions regarding a product, service, concept, idea etc. 	<ul style="list-style-type: none"> ▪ Work by consensus between informants ▪ Collect views from more than one informant at the same time ▪ Quick to identify important issues 	<ul style="list-style-type: none"> ▪ Same drawbacks as individual interviews ▪ Skilled facilitator needed ▪ May inhibit expression of minority views ▪ Views of the group cannot be generalised
<p>Questionnaire surveys</p> <ul style="list-style-type: none"> • Assess strengths and weaknesses in a quantitative way, e.g. estimating % of users satisfied with a given FAO product or service, can also assess participant perceptions, opinions and recommendations through textual inputs 	<ul style="list-style-type: none"> ▪ Powerful data collection tools: larger group of informants ▪ Objectivity: data collected in a standard and formal way ▪ Statistical data analysis: can examine trends and patterns quantitatively 	<ul style="list-style-type: none"> ▪ Designing good questionnaires is difficult ▪ Response rate often low ▪ Often no obligation to respond to questionnaire, hence sample is biased towards most opinionated

⁹ Extracts from the PIRES application outline a description of PEs, planned partnerships, allocated resources and planned outputs of the MTP, as well as the new Programme structure (reflected in the Monitoring and Assessment module, which also provides a log of implementation milestones of Programme outputs).

PART III: MANAGING AN AUTO EVALUATION

A. Selecting and Sequencing Data Collection Techniques

It is useful to start an auto-evaluation with a desk study of documented material, followed by a SWOT analysis with Programme staff (Diagram 3).

A desk study is a good way to start in order to:

- ▣ Establish what the original plan was, including indicators defined (e.g. extracts from PIRES MTP),
- ▣ Draw a list of all outputs produced over the evaluated period,
- ▣ Identify implementation constraints,
- ▣ Identify main implementation partners, and
- ▣ Assess outcomes.

During the initial period of background information gathering, a SWOT analysis is an effective way to engage with Programme staff to assess what critical issues could be evaluated. SWOT analysis findings can also be useful at a later stage to strategically address auto-evaluation recommendations (Diagram 3).

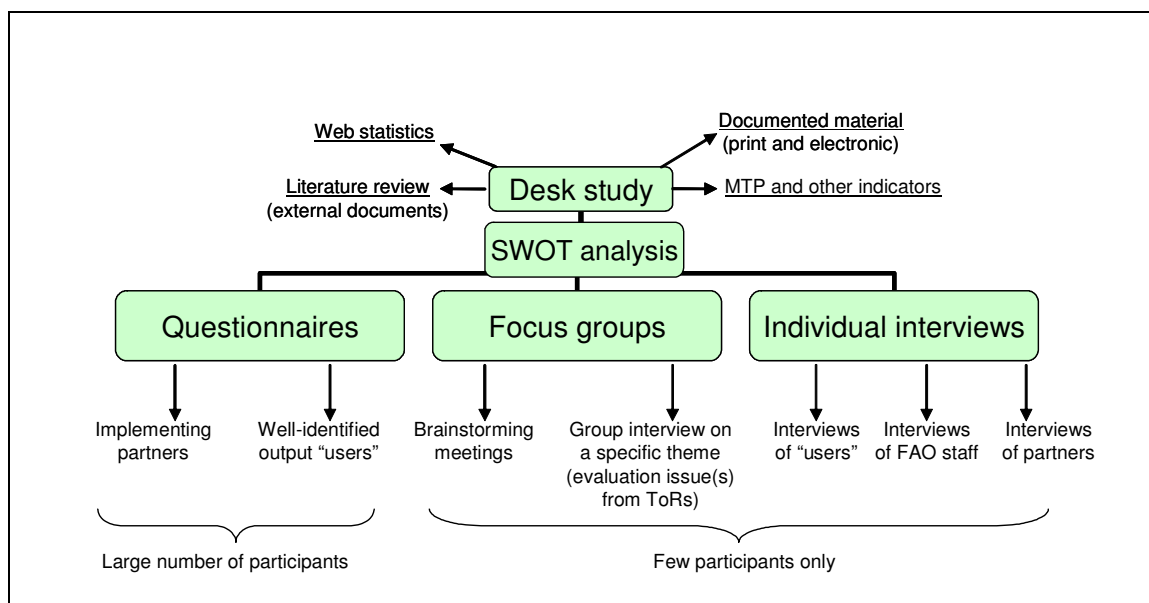
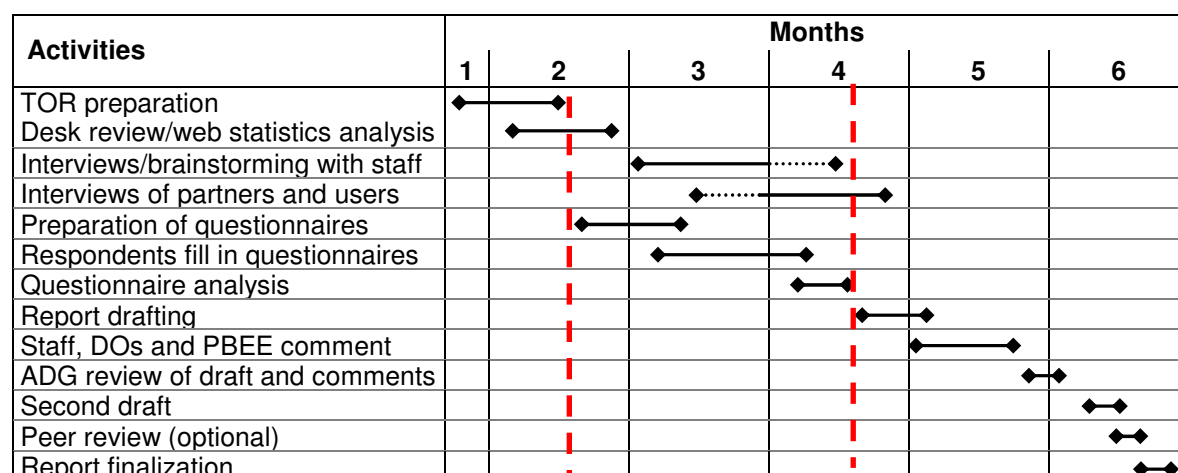


Diagram 3: An outline of sequencing common data collection techniques

While each auto-evaluation has its own timeline and choice of techniques, the most common way of sequencing data collection techniques is illustrated in Table 3.

Table 3: Example of an Auto-Evaluation Timeline



NOTE: The time period between the dashed lines reflects the most active phase of involvement of the external consultant in the scenario used above (approximately 2 months).

B. Following-Up on Auto-Evaluation Recommendations

In line with other types of evaluation employed in FAO, auto-evaluations also involve a Management Response mechanism. A format for Management Response to AE recommendations forming a basis for Management decisions is outlined in Table 4. As part of the Management Response mechanism, two key meetings are envisaged after the completion of the auto-evaluation, involving the AE manager, Programme Co-ordinator and staff. A meeting immediately after the AE should aim to outline a Management Response, whereas a follow-up meeting a year after the AE should revisit the Management Response.

Table 4: Management Response Format to AE Recommendations

Format for Management Response to Auto-evaluation Recommendations								
Recommendations	Further (RP or extra-budgetary) funding required (yes or no)	Acceptance by Management			Comment	Action to be taken		
		Accept	Partially Accept	Reject		Action	Timing	Unit Responsible
Recommendation 1								
Recommendation 2 etc.								

NOTE: The management response will be sent to the Chief of PBEE within one month after completion of the AE report.

ANNEXES

- Annex 1: Examples of Issues in PE Auto-Evaluations
- Annex 2: Outline for Terms of Reference
- Annex 3: Auto-Evaluation Report Outline
- Annex 4: Desk Study
- Annex 5: SWOT analysis
- Annex 6: Interviews and Focus Groups
- Annex 7: Questionnaires
- Annex 8: Web Statistics
- Annex 9: Bibliography

Annex 1: Examples of Issues in PE Auto-Evaluations

Table 5 shows only **generic examples** of what a list of evaluation issues can look like, **not a limitative or prescriptive list** and should be adjusted to the specific issues of each individual auto-evaluation. The categorization used below is also only one way to thematically structure evaluation issues. A recommended way to elaborate a list of evaluation issues is through engaging with and receiving feedback from Programme staff and partners. All AEs should include a description and assessment of achieved outcomes.

Table 5: Generic Issues in PE Auto-Evaluations

Design Issues
<ul style="list-style-type: none"> ▪ Are the PE strategy and hierarchy of objectives (outputs, outcomes, objectives) coherent & achievable? ▪ Are the resources allocated to the PE sufficient to deliver the outputs? ▪ Does FAO have a clear comparative advantage, mandate and priority on the subject matter?
Implementation and Process Issues
<ul style="list-style-type: none"> ▪ Are planned resources (human and financial) actually available to the PE well utilized? ▪ Is FAO working with the right partners and competencies on the subject matter? ▪ Are outputs produced at a reasonable cost and with accepted quality standards?
Output Issues
<ul style="list-style-type: none"> ▪ Which outputs did the PE produce during the evaluated period, and how does this list compare with planned outputs? ▪ How do biennial outputs contribute to their major outputs? ▪ Is there an effective dissemination strategy for FAO outputs?
Outcome Issues
<ul style="list-style-type: none"> ▪ What is the actual audience? What sort & number of users are reached by FAO products & services? ▪ What do they think of FAO outputs & what do they do with them? ▪ Are there any unplanned outcomes (positive or negative) resulting from the PE?
Objective-level Issues
<ul style="list-style-type: none"> ▪ What contributions to improved decision-making at international, national or sub-national levels can be documented or conjectured from existing evidence (e.g. in governments, donors, UN agencies, community organizations and NGOs)?
Cross-Sectoral Issues
<ul style="list-style-type: none"> ▪ How has the evaluated initiative contributed to the goals of the PAIAs it was planned to participate in? ▪ Was any progress made on gender mainstreaming, e.g. in the priority areas identified in the Gender Plan of Action (gender-segregated data, gender-sensitive communication strategies, equal access to natural resources and agricultural support systems, gender-sensitive policy and planning, etc.).

Annex 2: Outline for Terms of Reference

1. Background

Provides the context for the evaluation and should indicate as a minimum:

- a) a description of the PE as designed: objectives, planned major outputs; starting and ending dates, budget, main inputs;
- b) a brief overview of the history behind the PE;
- c) a description of major activities and outputs to date; and
- d) problems or emerging issues identified by management.

2. Purpose of the Auto-Evaluation

This section should briefly state why the evaluation is being held, and should have a particular focus, remembering that the reason for evaluation is to provide an input for future direction.

3. Scope of the Auto-Evaluation

The following is a standard set of issues to address (see also Part II, Section A). However, the scope of each AE should also be customized to specific concerns and questions of a given PE and additional unforeseen issues should be added to this list:

- a) Relevance of the PE to development priorities and needs of Member Nations;
- b) Clarity, logical consistency and realism of the PE design, including specification of inputs, outputs, outcomes and objectives, targets, identification of users and beneficiaries;
- c) Realism and clarity of external institutional relationships, and in the managerial and institutional framework for implementation;
- d) Efficiency of project implementation including: availability of funds and human resources as compared with budget; managerial and work efficiency; and implementation difficulties;
- e) Results, including a systematic assessment of outputs produced to date (both quantity and quality) and progress towards the realization of the PE outcomes and objective;
- f) Contributions to gender and social equity, in particular contributions to the Gender Plan of Action, and contributions to PAIAs;
- g) Prospects for sustaining the results by external users and partners after the termination of the PE;
- h) Overall effectiveness (including cost-effectiveness) of PEs; and
- i) A review of emerging issues of particular importance to management.

4. Roles in the Auto-Evaluation

This section should describe:

- a) who is responsible for overseeing and coordinating the evaluation;
- b) the composition and competencies of the evaluation team, including external consultant(s); and

- c) the composition and competencies of the peer review panel, if envisaged.

5. Methodology

This section should outline the main evaluation techniques to be used, including the information source (participants) and a rough time frame for the envisaged techniques. These may include, though not limited to the following:

- b) Desk study, including indicators to be measured;
- c) SWOT analysis and/or other group facilitation techniques;
- d) Interviews / focus group interviews
- e) Questionnaires; and
- f) Peer reviews, if envisaged.

6. Auto-Evaluation Outputs

The evaluation outputs section should describe the reporting stages in terms of procedure/debriefing arrangements (e.g. identifying staff/consultant involvement), duration of each stage, and tentative deadlines for delivering draft and final reports.

7. Budget

Based on the above and notably sections 4 and 5, elaborate a budget that includes staff and external consultant resources, i.e. the estimated staff time spent on managing the evaluation plus the financial resources required for hiring an external consultant and travel expenses, if required. Table 6 shows an example of a typical outline of expenditure items in auto-evaluations:

Table 6: Estimating Budget for an Auto-Evaluation

Expenditure Item	Time	Cost (\$ USD)
Internal Staff Costs:		
AE oversight	X weeks @ _ USD per day	
AE manager	X weeks @ _ USD per day	
Programme staff	X weeks @ _ USD per day	
Other human resources	X weeks @ _ USD per day	
Sub-total (Internal Staff Costs)		
External Consultant Costs:		
Senior consultant	X days (WAE) @ _ USD per day	
Junior consultant	X days (WAE) @ _ USD per day	
Travel expenses	times @ _ USD	
DSA (Rome)	X days @ _ USD per day	
Sub-total (External Consultant Costs)		
TOTAL COSTS		

NOTE: The estimated budget should also include other expenditure items of substantial cost, if used, such as DSA for a peer review group.

Annex 3: Auto-Evaluation Report Outline

Following is a suggested general structure of an auto-evaluation report.

I. Executive Summary (Main Findings and Recommendations)

II. Introduction

III. Background and Context

IV. Relevance to Priorities and Needs of Member Nations

V. Assessment of Programme Entity Design

- A. Clarity, consistency and realism of the PE design (including inputs, outputs, outcomes and objectives, users and beneficiaries)
- B. Realism and clarity of external institutional relationships, and in the institutional framework for implementation

VI. Assessment of PE Implementation, Processes, Efficiency and Management

- A. Financial and human resource management
- B. Activities undertaken and outputs produced
- C. Partnerships and collaborative processes

VII. Assessment of Results and Effectiveness

- A. Audience of the PE and documented outcomes
- B. Progress towards the realization of the objective
- C. Achievements in terms of gender and social equity
- D. Cost-effectiveness
- E. Major factors affecting the project results

VIII. Conclusions and Recommendations

IV Lessons Learned

Annexes

- 1. Terms of reference
- 2. Key persons met or interviewed
- 3. Documents consulted by the mission

When preparing the auto-evaluation report, take into account the following:

- ▣ The report need not exceed 30 pages, and should display a critical and analytical outlook.
- ▣ The report should be structured on the basis of the evaluation issues outlined in the TOR so that the reader can compare the AE report with its TOR.
- ▣ If a team of evaluators is involved, all of them should participate in report writing. The evaluators should seek consensus and agree on their findings, conclusions and recommendations¹⁰.
- ▣ The use of summary tables and graphs for the presentation of quantitative findings, where possible, is recommended.
- ▣ Conversely, detailed client stories, including the use of verbatim quotes (signed quotes require asking for author consent to quote by name), are very good tools to present the best (or worst) that a particular programme has to offer in a lively, telling manner.
- ▣ Conclusions involve subjective judgment calls, e.g. on whether or not specific activities are worth the expense and efforts invested in them. The report writer should clearly distinguish respondent perceptions collected, from factual descriptions through the use of phrases such as “the evaluators conclude that...” or “on the basis of partner feedback....”
- ▣ Given that FAO has to operate within limited resources, there should be a set of recommendations not requiring increased resources, in addition to any that call for additional resources. Recommendations (5 to a maximum of 15) should be precise, realistic, creative, actor-specific, and if many, thematically structured.
- ▣ Any proposal for a new or revised PE should provide a rough sketch of its design that can be used as an input in the next MTP preparation.

¹⁰ If strong disagreements surface that cannot be resolved, those holding the minority view may consider writing a dissenting opinion that will be annexed to the main report (seldom-used measure).

Annex 4: Desk Study

A desk study is a common starting point of an auto-evaluation and typically involves:

- ▣ Objectives, planned outputs, indicators and partnerships related to the PE(s), as well as previous Programme Structure mapping into the New Programme Structure (sourced from PIREs and using both MTP and the Monitoring and Assessment modules);
- ▣ Documented programme outputs – a compilation and summary of all relevant documents, such as programme documents, workplans, back-to-office reports, meeting minutes; and
- ▣ Literature reviews, by analysing documents *external* to a programme, i.e. publications (either print or electronic) produced by the technical or mass media, by other programmes or organizations – in order to gauge which other publications or media are quoting FAO (i.e. to gauge the FAO audience).

Auto-evaluators can choose additional quantitative and qualitative indicators (e.g. Table 7) to those defined in the MTP in order to compare planned with achieved outputs. Proxy indicators (measuring an output indirectly) can be used when necessary, given that the cost, complexity and timeliness of data collection may prevent outputs being measured directly. Indicators based on relevant disaggregated data should be used when available – this means that data is segmented according to location, gender, income level, social group and any other relevant characteristic. However, note that indicators only reflect change over time and are suitable for trend analysis, but do not explain why change is happening.

Literature reviews are useful to gauge main issues and opinions being debated by “opinion makers” with whom FAO interacts (or wishes to interact), assess whether FAO is making a unique contribution or is merely repeating the work of others, and collect references to FAO work in publications of others, so as to gauge the FAO audience. The later is conducted by analysing selected citations (e.g. title of a flagship publication) using a specific search engine (e.g. Google, Google Scholar, Elsevier’s Scopus). Note that this technique (citation analysis) has certain limitations (bound to Internet publications and limited to “specialization” of search engine used)

Table 7: Indicators, targets and achievements

Indicators	Targets	Achievements		
		2001	2002	2003
A. List of countries collecting data using methodology XXX	20	5	10	15
B. Number of national staff exposed to new approach through workshop	2,000	500	1,800	2,500
C. Examples of lessons learned and agreed at international conference	Consensus about issue YYY	Consensus achieved	n.a.	n.a.

NOTE: Examples A. and B. reflect quantitative indicators, whereas C. is an example of a qualitative indicator.

Annex 5: SWOT Analysis

SWOT analysis is a facilitated exercise, conducted during the development of the TOR, whereby Programme staff identify their unit's strengths and weaknesses, as well as the opportunities and threats they face. A SWOT exercise can be facilitated by external facilitators (e.g. PBEE Auto-Evaluation Advisors) and presents an opportunity for building team consensus, engaging Programme staff with the AE process and fostering ownership of the process, as well as gaining constructive staff input regarding PE issues.

The first step in a SWOT session is to ask participants to fill in cards with the strengths, weaknesses, opportunities and threats they perceive as important. These are then placed on corresponding pin-boards. Each idea should then be discussed and agreed upon by the whole group. Only those cards on which the whole group agrees would normally remain on the pin-boards, as a reflection of group consensus. Duplicates or very similar ideas should be rephrased as one card. Alternatively, the facilitator may write ideas up on flipcharts. Finally, the agreed list can be ranked according to importance. Following are the elements of a SWOT exercise:

- ▣ **Strengths:** qualities, assets or strong points inherent to or associated with the group, which may enable the group to fulfil a useful and/or expanded role.
- ▣ **Weaknesses:** traits that have a direct negative impact on the group's work performance and hinder its ability to perform an appreciated role and remain relevant. To the extent possible, weaknesses should be eliminated or minimized.
- ▣ **Opportunities:** external factors, circumstances or trends which can potentially and favourably affect a group's ability to operate, and can lead to a positive development of the group's role, if they are exploited.
- ▣ **Threats:** external circumstances, competition or risks, which might unfavourably influence group activities or relevance, and/or might result in an undesirable development. To the extent possible, threats have to be overcome or mitigated.

Beside a SWOT exercise, other creative facilitated group work can be used. For example, for assessing strengths and weaknesses, participants can be asked to write down on separate pieces of paper things of their concern (problems) and things that they would like to change (wishes). After gathering all participant input, the facilitator writes up a so called "wish" list and a "bug" list on a flipchart. All the items/issues on both lists are discussed and duplicates or very similar ideas are rephrased as one item on the list. After the two lists have been agreed on, items in both the "wish" list and the "bug" list are ranked according to importance. Participants should be encouraged to be open and creative in their feedback and to express even what may appear as a trivial issue within their working environment.

Annex 6: Interviews and Focus Groups

Interviews are most useful when one wants to collect complex, qualitative information, such as perceptions or client stories, i.e. narratives of how a specific user has applied or disseminated (or not) FAO products and services.

- ▣ Semi-structured interviews use a checklist of open-ended questions, but the interviewer is free to go over and beyond the list. Responses to questions provide the evaluators with quotations, which compose the raw data.
- ▣ Individuals to be interviewed should be contacted as early as possible to fix appointments in advance. Programme staff should be first interviewed, followed by external informants.
- ▣ One should allow 2-3 weeks minimum for the development, circulation among concerned staff, testing, finalization and launch of questionnaires and a month for getting back an adequate number of completed questionnaires¹¹.

In terms of timing, interviews can be used:

- ▣ *Before* designing quantitative, structured questionnaires, as interviews can help prepare a questionnaire by identifying issues and typical answers to evaluation questions; or
- ▣ *After* results of more standardized measures are analyzed, in order to gain further insight into interesting or unexpected findings, especially as follow-up individual interviews (e.g. in-person, by phone or email) for those questionnaire respondents that appear particularly knowledgeable.

Focus groups involve a group of individuals discussing a set topic, brought together by a skilled and experienced facilitator who uses the group and its interaction as a way to gain information about a specific issue. Key principles and factors to be observed in organizing focus groups are:

- ▣ Focus group interviews should not be rushed: the facilitator should first establish contact, describe the evaluation purpose and ground rules before asking the first question;
- ▣ Groups are typically composed of 7 to 10 people per focus group;
- ▣ Validity of individual ideas can be checked in the group;
- ▣ Focus groups can provide data more quickly and at lower cost than individual interviews;
- ▣ Focus group interviews may inhibit the expression of minority views among the group;
- ▣ Useful to study interactions, consensus or disagreements between participants; and
- ▣ In some cases, each stakeholder group needs to be interviewed separately¹².

¹¹ Reminder letters or emails are usually necessary to elicit good response rates.

Annex 7: Questionnaires

Questionnaires distributed through the mail, email and web media are a powerful tool for collecting quantitative and qualitative information on Programme performance and outputs from a user perspective. However, it is most appropriate to be used when the potential sample of respondents is large, easily identified and accessible. A well-designed survey requires questionnaire design skills. PBEE offers questionnaire design training on request, to auto-evaluation teams. Different types of surveying methods are compared in Table 8.

Table 8: Advantages and disadvantages of surveying methods

Methods	Advantages	Disadvantages
Face-to-face Interviews	<ul style="list-style-type: none"> ▪ Rich and in-depth source of information from key informants ▪ Longer interviews are better tolerated than longer questionnaires ▪ Interviewers can ask for clarifications and elaborations 	<ul style="list-style-type: none"> ▪ Cost per interview higher than for any other method, and extremely high for most AEs as target population is dispersed in member countries ▪ Possible interviewer bias
Telephone Interviews	<ul style="list-style-type: none"> ▪ Fast ▪ Less expensive than face-to-face ▪ Long interviews are sometimes tolerated (less than face-to-face) ▪ Interviewers can ask for clarifications and elaborations 	<ul style="list-style-type: none"> ▪ Requires the phone numbers of a sample of respondents ▪ Less control of the interview environment for the respondent than in face-to-face situations ▪ Possible interviewer bias
Feedback forms *	<ul style="list-style-type: none"> ▪ Very fast ▪ The facilitator can motivate participants to respond, hence good representativeness of results 	<ul style="list-style-type: none"> ▪ Usually limited to the evaluation of one particular event, thus rarely applicable to entire PEs
Mail Questionnaires	<ul style="list-style-type: none"> ▪ Less expensive than interviews ▪ Less intrusive than interviews ▪ The respondents answer at their leisure 	<ul style="list-style-type: none"> ▪ Quite slow to get the results (several weeks) ▪ Self-selected, and thus biased sample ▪ Problems with mail delivery
Email Questionnaires	<ul style="list-style-type: none"> ▪ Very economical and fast ▪ In developing countries, email access may be better than web access ▪ Response rates usually higher (novelty of the method, no need to mail back the questionnaire) 	<ul style="list-style-type: none"> ▪ Many people dislike unsolicited email even more than unsolicited regular mail ▪ Limited to target populations with good email connections ▪ Self-selected, and thus biased sample
Web Questionnaires	<ul style="list-style-type: none"> ▪ Low data entry cost ▪ Can automate skip patterns based on earlier answers ▪ Response rates usually higher (novelty of the method, no need to mail back the questionnaire) ▪ Can be combined with an email invitation to take a Web survey 	<ul style="list-style-type: none"> ▪ Limited to target populations with good Internet access ▪ Self-selected, and thus biased sample ▪ Security must be well thought out to avoid anyone browsing that Web page to answer ▪ Some may answer several times ▪ Identifying respondents

* For example, brief questionnaires distributed at the end of a training or workshop.

¹² For example, stakeholder groups should be separately interviewed when there is strong conflict of interest or if participants have a diverse background between the groups.

Annex 8: Web Statistics

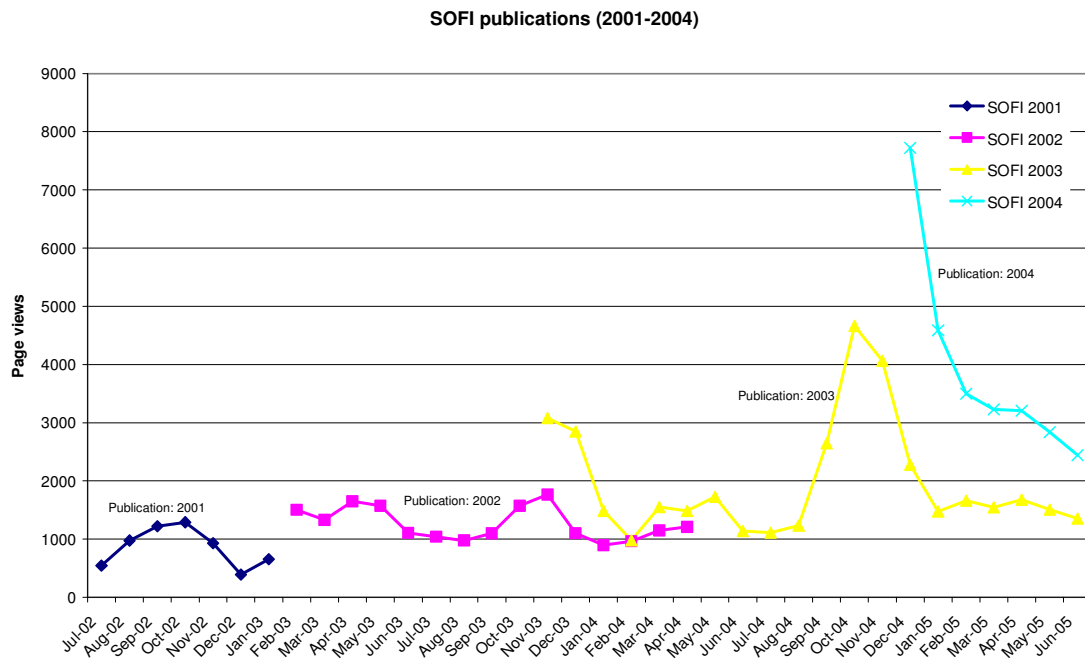
Web statistics allow a crude estimate of the number of users at one site, which pages are most frequently viewed, how they had arrived at the site (referring sites) etc. However, the system can supply limited information on the characteristics of web site users and has a number of other limitations (e.g. difficulty tracing geographical/institutional origin of user or repeat use of a website, limited interpretation of hits and page views etc.). Without some form of feedback from users on their experience in the use of information systems, there can be no sound basis for redesign efforts (e.g. ask users to register their e-mail addresses, gender, nationality, location, institutional affiliation and areas of interest in order to gain access to certain parts of the site, or a user questionnaire).

As part of information systems and web site related work, some FAO technical departments follow web traffic of their respective division/service level websites. Furthermore, The Knowledge and Communication Department (KCEW) provide regular FAO website statistics, currently using the AWStats¹³ open source application.

Some PEs may have well-established flagship publications and assessing publication readership through the use of web statistics, in these cases, may be particularly useful for analysing trends. This usually requires:

- ▣ Extraction of selected monthly data (e.g. Internet visits of a key publication) from KCEW web traffic reports (Annex 9), for a given period, into an Excel file¹⁴; and
- ▣ Graphing the monthly data to show trends over time (see Figure 1).

Figure 1: An example of trend analysis of a flagship publication using web traffic statistics



¹³ GILW previously used WebTrends, a commercial package for web traffic analysis and switched to a more cost-effective AWStats package by the end of 2005.

¹⁴ WebTrends/AWStats data is currently limited to “manual” extraction into Excel.

Annex 9: Bibliography

Terminology:

The most widely accepted evaluation terminology can be found in the OECD glossary of evaluation terms, in English, French and Spanish:
<http://www.oecd.org/dataoecd/29/21/2754804.pdf>

Planning for evaluation:

The American Evaluators Association and the Michigan University maintain a number of checklists for evaluation management using various evaluation models. A bit too detailed for non-specialists but a useful source of ideas and tips nonetheless.
<http://www.wmich.edu/evalctr/checklists/checklistmenu.htm>

Indicators:

A short primer on indicator theory from UNDP:
http://stone.undp.org/undpweb/eo/evalnet/docstore3/yellowbook/documents/key_indicator_s.pdf

In the 90's the World Bank invested quite a lot of efforts into its "Performance Monitoring Indicators", in effect a list of standard indicators for all development sectors (energy, micro-finance etc.). Their handbook discussing indicators is here: http://www-wds.worldbank.org/servlet/WDS_IBank_Servlet?pcont=details&eid=000009265_3961219_094954

An EU paper on common questions with criteria and indicators:
http://europa.eu.int/comm/agriculture/rur/eval/evalquest/a_en.pdf

The US Environment Protection Agency has given some thoughts to the evaluation of indicators: <http://www.epa.gov/indicators/qformat.htm>

Surveys:

A very good and short tutorial on survey design, by David S. Walonick:
<http://www.statpac.com/surveys/surveys.doc> or <http://www.statpac.com/surveys/> (Statpac is selling a survey software as well).

A useful list of response scales (Fair / Unfair; Agree / Disagree, etc.) at
<http://dataguru.org/ref/survey/responseoptions.asp>

A concise guide to questionnaire development, developed for surveying the usability of information technology projects: <http://atwww.hhi.de/USINACTS/tutorial/quest.html>

Sites on statistical analysis:
<http://www.statsoft.com/textbook/stathome.html>

Qualitative approaches - semi-structured interviews and focus groups:

A good primer to qualitative evaluation techniques can be found on the Cyfer.net website of the University of Arizona (apparently being restructured), in particular "[Using Focus Groups for Evaluation](#)", by Mary Marczak & Meg Sewell, and "[The Use of Qualitative Interviews in Evaluation](#)", by Meg Sewell.

Evaluating programme outcomes:

The IDRC evaluation office has developed a detailed methodology called "outcome mapping" to help programme evaluators and staff review their effectiveness.

http://web.idrc.ca/ev.php?url_id=26586&url_do=do_topic&url_section=201&reload=1060162955

The UNDP Evaluation Handbook, largely relying on outcome mapping:

http://stone.undp.org/undpweb/eo/evalnet/docstore3/yellowbook/documents/full_draft.pdf

Web statistics and evaluation of websites:

FAO web traffic reports (pre-2005 data, prepared by KCEW):

<http://waicent003/OldReports/>

FAO web traffic reports (post-2005 data, prepared by KCEW):

<http://awstats.fao.org/directory>

Hope N. Tillman: Evaluating Quality on the Net. Babson College, Massachusetts:

<http://www.hopetillman.com/findqual.html#my>

Smith, Alastair G. "Testing the Surf: Criteria for Evaluating Internet Information Resources." The Public-Access Computer Systems Review 8, no. 3 (1997):

<http://info.lib.uh.edu/pr/v8/n3/smit8n3.html>

Peer reviews

OECD-DAC has a rich collection of evaluation guidelines, including: Peer Review: a Tool for Co-Operation and Change - an Analysis of an OECD Working Method by Fabrizio Pagani: <http://www.oecd.org/dataoecd/33/16/1955285.pdf>

Same document in French: L'Examen par les Pairs : un Instrument de Coopération et de Changement. <http://www.oecd.org/dataoecd/33/18/1955301.pdf>



Food and Agriculture Organization of the United Nations
Evaluation Service (PBEE)
Viale delle Terme di Caracalla
00100 Rome, Italy
<http://www.fao.org/pbe/pbee/>
Comments should be sent to: evaluation@fao.org